

AMENDMENT(S) TO THE CLAIMS:

The following listing of claims will replace all prior versions, and listings, of claims on the application. All claims are set forth below with one of the following annotations.

- (Original): Claim filed with the application.
- (Currently amended): Claim being amended in the current amendment paper.
- (Canceled): Claim cancelled or deleted from the application. No claim text is shown.
- (Withdrawn): Claim still in the application, but in a non-elected status.
- (New): Claim being added in the current amendment paper.
- (Previously presented): Claim added or amended in an earlier amendment paper.
- (Not entered): Claim presented in a previous amendment, but not entered or whose entry status unknown. No claim text is shown.

1. (Previously presented) An interactive spatialized audiovisual system for linking a plurality of remote user terminals, the system comprising:

a processing system coupled to a network; and;

an associated user database coupled to or part of the processing system, the user database including user data, including user status information for each corresponding user,

wherein the processing system is configured to:

receive a plurality of audio streams and associated locating data from the remote user terminals each audio stream corresponding to a user as a source of audio, the locating data for virtually locating the users relative to one another within a virtual user environment;

select at least one group of at least one of the plurality of audio streams based on status information in the user database, each group corresponding to one of the users,

combine at least some of the plurality of audio streams to form a combined stream, and

send to each of at least one of the remote user terminals via the network the respective selected group of audio streams and associated locating data corresponding to the user of the remote user terminal, and

send to the at least one of the remote user terminals via the network a function of the combined stream; the function possibly user dependent,

wherein, a particular remote user terminal coupled to the network and corresponding to a particular user is configured to:

receive the sent group of audio streams and the function of the combined stream;

display a visual representation of the virtual user environment, including representations of at least some of the users; and

convert the audio streams of the group corresponding to the particular user and the function of the combined stream to a pair of audio headphones signals including binauralized reverberant signals generated according to the combined stream,

wherein the converting includes spatializing the audio streams of the group such that the particular user, listening to the headphone signals over headphones, has the sensation that the audio streams of the group are emanating from their respective user locations in the virtual user environment relative to the location of the particular user,

wherein the spatializing includes HRTF processing to take into account the orientation and location of the particular user in the virtual user environment, and to take into account direct sounds and early echoes, and reverberation according to a non-spatial combination of audio streams.

2. (Previously presented) A system according to claim 1, wherein the processing system is further configured to carry out at least part of the converting of the audio streams of each group of audio streams.
3. (Previously presented) A system according to claim 1, wherein the processing system is further configured to spatialize a reverberantly processed version of the combined stream such that the particular user listening to the headphone signals over headphones perceives a reverberant background formed from the combined stream.
4. (Previously presented). A system according to claim 1, wherein at least part of the spatializing occurs at the processing system such that the sending of a group of audio streams to the particular user is in at least a partially spatialized form.
5. (Previously presented) A system according to claim 1, wherein the user status information used in the selecting the group to send to a particular user includes one or more of user location data which serves to locate the particular user in the virtual environment, user orientation data which serves to orientate the particular user both with respect to the other users and to the virtual environment, user listening status information for the particular user, and/or user talking status information for the particular user.
6. (Previously presented) A system according to claim 5, wherein the user listening status information used in the selecting the group to send to a particular user is arranged to allow the particular user to listen to other selected users or groups of users in the environment.
7. (Previously presented) A system according to claim 5, wherein the selecting of a group of a particular number of one or more audio streams to send to a particular user includes selecting the closest particular number of other users within the virtual environment.
8. (Previously presented) A system according to claim 5, wherein the selecting of a group of a particular number of one or more audio streams to send to a particular user

includes selecting the loudest particular number of sources based on at least one of the the amplitude of the audio in the audio stream and/or the distance of the source of the audio stream from the particular user within the virtual environment.

9. (Previously presented) A system according to claim 5, wherein the selecting of a group of a particular number of one or more audio streams to send to a particular user includes selecting according to input from the particular remote user terminal or from other remote user terminals.
10. (Previously presented) A system according to claim 5, wherein the selecting of a group of a particular number of one or more audio streams to send to a particular user includes selecting according to input from a moderator able to control the talking and listening status of the users.
11. (Previously presented) A system according to claim 5, wherein the virtual environment has one or more geographical or topological features arranged to affect the listening capability of users in the virtual environment and wherein the selecting of a group of a particular number of one or more audio streams to send to a particular user includes selecting based on the geographical or topological features.
12. (Currently amended) A system according to claim 11, wherein the one or more features include one or more barriers defining ~~one~~ two or more chat rooms, wherein at least some of the audio streams in a particular room are summed and spatialized and wherein reverberation processing is performed to achieve a background reverberation effect characteristic of that particular room.
13. (Previously presented) A system according to claim 11, wherein the processing system is further configured to combine the audio streams in adjoining rooms or areas that have an entrance or exit in the virtual environment of the particular room or area where the particular user is located so as to create an adjoining room signal located at the entrance or exit, wherein the adjoining room signal is representative of the

combined noise emanating from the room or area adjoining the particular room or area at the entrance or exit .

14. (Previously presented) A system according to claim 13, wherein the processing system is further configured to generate for the particular room or area an in-room signal representative of the background due to audio streams emanating from users within the particular room or area and optionally including a weighted reverberant version of any adjoining room signal emanating from users in any room adjoining the particular room or area, wherein said generating includes combining the audio streams emanating from users within the particular room or area and optionally a weighted reverberant version of the combined signals emanating from users in the room or rooms or area or areas adjoining the particular room or area .
15. (Previously presented) A system according to claim 13, wherein the processing system is further configured to generate for the particular room or area a room signal representative of a reverberant version of the adjoining room signal formed as a combining of the signals emanating from users in the room adjoining the particular room or area.
- 16.–21. (Cancelled).
22. (Previously presented) A method of operating a processing system in communication with a plurality of remote user terminals comprising:
 - receiving a plurality of audio streams and associated locating data, each audio stream corresponding to a user as a source of audio, the locating data capable of virtually locating the audio sources relative to one another within a virtual user environment;
 - determining user status data for one or more of the users;
 - selecting at least one group of at least one of the plurality of audio streams based on the user status data, each group corresponding to one of the users;

combining at least some of the plurality of audio streams to form a combined stream; and

transmitting to each of at least one of the remote user terminals the respective selected group of audio streams and associated locating data corresponding to the user of the remote user terminal ; and

transmitting to the at least one of the remote user terminals a function of the combined stream; the function possibly user dependent,

wherein, a particular remote user terminal corresponding to a particular user is configured to:

receive the transmitted group of audio streams and the function of the combined stream;

display a visual representation of the environment, including representations of at least some of the users; and

convert the audio streams of the group corresponding to the particular user and the function of the combined stream to a pair of audio headphones signals including binauralized reverberant signals generated according to the combined stream,

wherein the converting includes spatializing the audio streams of the group such that the particular user, listening the headphone signals over headphones, has the sensation that the audio streams of the group are emanating from their respective user locations in the virtual user environment relative to the location of the particular user, and

wherein the spatializing includes HRTF processing to take into account the orientation and location of the particular user in the virtual user environment, and to take into account the orientation and location of the particular user in the virtual user environment, and to take into account direct sounds and early echoes, and reverberation according to a non-spatial combination of audio streams,

wherein the combined stream is spatialized either before or after transmitting so as to provide a background audio effect within the virtual environment.

23. (Previously presented) A method according to claim 22, further comprising repeating the determining and the selecting steps to allow for the case that the user status data is altered.

24. (Previously presented) A method according to claim 22, wherein the virtual environment has one or more barriers defining two or more rooms or areas, the method further comprising:

combining the audio streams in adjoining rooms or areas that have an entrance or exit in the virtual environment of the particular room or area where the particular user is located so as to create an adjoining room signal located at the entrance or exit, wherein the adjoining room signal is representative of the combined noise emanating from the room or area adjoining the particular room or area at the entrance or exit .

25. (Previously presented) A method according to claim 24, further comprising:

generating for the particular room or area in the virtual user environment an in-room signal representative of the background due to audio streams emanating from users within the particular room or area and optionally including a weighted reverberant version of the adjoining room signal emanating from users in any room adjoining the particular room or area, wherein said generating includes combining the audio streams emanating from users within the particular room or area and optionally a weighted reverberant version of the combined signals emanating from users in the room or rooms or area or areas adjoining the particular room or area.

26. (Previously presented) A method according to claim 24, wherein the combining the audio streams in adjoining rooms or areas includes generating for the particular room or area a room signal representative of a reverberant version of combination of the signals emanating from users in the room adjoining the particular room or area.

27. (Canceled).

28. (Currently amended) A user terminal for participating in a spatialized conversation in a network environment, the user terminal coupled to a computer network capable of streaming audio streams and associated spatialization information to the user terminal, the user terminal comprising:

a rendering system configured to:

accept a selected group of audio streams selected from a plurality of audio streams, each stream corresponding to a user at a user location in a virtual user environment ~~to the user terminal~~;

accept associated locating data for virtually locating the users associated with the group's audio streams relative to one another within the virtual user environment,

accept a function of a combined stream formed by combining at least some of the plurality of audio streams; and

convert the audio streams of the group and the function of the combined stream to a pair of audio headphones signals including binauralized reverberant signals generated according to the combined stream,

wherein the converting includes spatializing the audio ~~stream~~ streams of the group such that the particular user, listening to the headphone signals over headphones, has the sensation that the audio streams of the group are emanating from their respective user locations in the virtual user environment relative to the location of the particular user; and

a user interface including a display configured to present a visual representation of the virtual user environment, including representations of at least some of the users,

wherein the spatializing by the rendering system includes HRTF processing to take into account the orientation and location of the particular user in the virtual user

environment, and to take into account direct sounds and early echoes, and reverberation according to a non-spatial combination of audio streams.

29. (Currently amended) A user terminal according to claim 28, wherein the virtual environment has one or more barriers defining two or more rooms or areas, wherein the particular user is located in the virtual environment in a particular room or area that ~~includes~~ has an adjoining room or area that has an entrance or exit ~~in the virtual user environment~~ to the particular room or area, wherein the rendering system is further configured to accept an adjoining room signal located at the entrance or exit, and wherein the adjoining room signal is representative of the combined noise emanating from the room or area adjoining the particular room or area at the entrance or exit .
30. (Currently amended) A user terminal according to claim 28, wherein the virtual) A user terminal according to claim 29, wherein the rendering system is configured to receive an in-room signal representative of the background due to audio streams emanating from users within the particular room or area and optionally including a weighted reverberant versions of any adjoining room signal emanating from users in any room or area adjoining the particular room or area, wherein the in-room signal is formed by a process that includes combining the audio streams emanating from users within the particular room or area and optionally a weighted reverberant version of the combined signals emanating from users in the room or area or rooms or ~~rooms~~ areas adjoining the particular room or area .
31. (Previously presented) A user terminal according to claim 29, wherein the adjoining room signal is formed by a process that includes combining the audio streams in adjoining rooms or areas including generating for the particular room or area a room signal representative of a reverberant version of combination of the signals emanating from users in the room adjoining the particular room or area .
32. (Canceled).

33. (Previously presented) A computer-readable medium having stored thereon executable instructions that when executed by one or more processors of a processing system in communication with a plurality of remote user terminals, cause implementing a method comprising:

receiving a plurality of audio streams and associated locating data, each audio stream corresponding to a user as a source of audio, the locating data capable of virtually locating the audio sources relative to one another within a virtual user environment;

determining user status data for one or more of the users;

selecting at least one group of at least one of the plurality of audio streams based on the user status data, each group corresponding to one of the users;

combining at least some of the plurality of audio streams to form a combined stream; and

transmitting to each of at least one of the remote user terminals the respective selected group of audio streams and associated locating data corresponding to the user of the remote user terminal ; and

transmitting to the at least one of the remote user terminals a function of the combined stream; the function possibly user dependent,

wherein, a particular remote user terminal corresponding to a particular user is configured to:

receive the transmitted group of audio streams and the function of the combined stream;

display a visual representation of the environment, including representations of at least some of the users; and

convert the selected audio streams of the group corresponding to the particular user and the function of the combined stream to a pair of audio headphones

signals, including binauralized reverberant signals generated according to the combined stream,

wherein the converting includes spatializing the audio streams of the group such that the particular user, listening the headphone signals over headphones, has the sensation that the audio streams of the group are emanating from their respective user locations in the virtual user environment relative to the location of the particular user, and

wherein the spatializing includes HRTF processing to take into account the orientation and location of the particular user in the virtual user environment, and to take into account direct sounds and early echoes, and reverberation according to a non-spatial combination of audio streams,

wherein the combined stream is spatialized either before or after transmitting so as to provide a background audio effect within the virtual environment.

34. (Previously presented) A computer-readable medium according to claim 33, wherein the virtual environment has one or more barriers defining two or more rooms or areas, wherein the method further comprises combining the audio streams in adjoining rooms or areas that have an entrance or exit in the virtual user environment of the particular room or area where the particular user is located so as to create an adjoining room signal located at the entrance or exit, and wherein the adjoining room signal is representative of the combined noise emanating from the room or area adjoining the particular room or area at the entrance or exit.

35. (Currently amended) A method of operating a particular user terminal that is part of an interactive spatialized audio facility including a networked computer and a plurality of user terminals linked to the networked computer, including the particular user terminal, the method comprising:

transmitting from the particular user terminal to the networked computer an audio stream generated by a particular user and associated locating data capable of virtually locating the source of the audio stream generated by the user within a virtual

environment, such that the networked computer can select groups of audio streams corresponding to each user, selectively combine at least some of the audio streams, for each group select associated locating data for the sources of the audio streams in the group, wherein the selecting is according to user status data available at the networked computer;

receiving at the particular user terminal a particular selected group of a plurality of audio streams selected on the basis of the user status data for the particular user, and further receiving associated locating data for virtually locating the sources of the group's audio streams relative to one another within a virtual user environment;

receiving at the particular user terminal a function of a combined audio stream formed by combining at least some of the plurality of audio streams corresponding to the users;

generating at the particular user terminal visual representations of the sources of the audio streams to indicate virtual locations of the sources in the virtual user environment, and

converting the selected group of audio streams and the function of the combined stream to a pair of audio headphones signals including binauralized reverberant signals generated according to the combined stream,

wherein the converting includes spatializing the audio ~~stream~~ streams of the group such that the particular user, listening to the headphone signals over headphones, has the sensation that the audio streams of the selected group are emanating from their respective user locations in the virtual user environment relative to the location of the particular user, and ,

wherein the spatializing includes HRTF processing to take into account the orientation and location of the particular user in the virtual user environment, and to take into account direct sounds and early echoes, and reverberation according to a non-spatial combination of audio streams.

36. (Previously presented) A method according to claim 35, wherein the virtual environment has one or more barriers defining two or more rooms or areas, and wherein the particular user is located in the virtual environment in a particular room or area that includes an adjoining room that has an entrance or exit in the virtual user environment, the method further comprising receiving at the user terminal an adjoining room signal representative of the combined noise emanating from the room or area adjoining the particular room or area and located at the entrance or exit between the particular room or area and the room or area adjoining the particular room or area .

37. (Currently amended) A method according to claim 36, further comprising:

receiving an in-room signal representative of the background due to audio streams emanating from users within the particular room or area and optionally including a weighted reverberant versions of any adjoining room signal emanating from users in any room or area adjoining the particular room or area, wherein the in-room signal is formed by a process that includes combining the audio streams of the users emanating from users within the particular room or area and optionally a weighted reverberant version of the combined signals emanating from users in the room or area or rooms or ~~rooms~~ areas adjoining the particular room or area

38. (Currently amended) A computer-readable medium having stored thereon executable instructions that when executed by a processor in a particular user terminal, cause carrying out of a method of operating the a particular user terminal, the user terminal being part of an interactive spatialized audio facility including a networked computer and a plurality of user terminals linked to the networked computer, including the particular user terminal, , the method comprising:

transmitting from the particular user terminal to the networked computer an audio stream generated by a particular user and associated locating data capable of virtually locating the source of the audio stream generated by the user within a virtual environment, such that the networked computer can select groups of audio streams

corresponding to each user, selectively combine at least some of the audio streams, for each group select associated locating data for the sources of the audio streams in the group, wherein the selecting is according to user status data available at the networked computer;

receiving at the particular user terminal a particular selected group of a plurality of audio streams selected on the basis of the user status data for the particular user, and further receiving associated locating data for virtually locating the sources of the group's audio streams relative to one another within a virtual user environment;

receiving at the particular user terminal a function of a combined audio stream formed by combining at least some of the plurality of audio streams corresponding to the users;

generating at the particular user terminal visual representations of the sources of the audio streams to indicate virtual locations of the sources in the virtual user environment, and

converting the selected group of audio streams and the function of the combined stream to a pair of audio headphones signals including binauralized reverberant signals generated according to the combined stream,

wherein the converting includes spatializing the audio ~~stream~~ streams of the group such that the particular user, listening to the headphone signals over headphones, has the sensation that the audio streams of the selected group are emanating from their respective user locations in the virtual user environment relative to the location of the particular user, and ,

wherein the spatializing includes HRTF processing to take into account the orientation and location of the particular user in the virtual user environment, and to take into account direct sounds and early echoes, and reverberation according to a non-spatial combination of audio streams.

39. (Currently amended) A system according to claim 5, wherein the selecting of a group of a particular number of one or more audio streams to send to a particular user includes selecting according to one or more of the criteria including:

selecting a number M of the audio streams corresponding to the M closest users from the total number N of audio streams,

selecting the number M loudest audio streams based on the amplitude of the signal of the audio stream and/or the distance of the user corresponding to the audio stream from the particular user,

selecting according to a user-driven selection process determined by the particular user or one ~~to~~ or more other users,

selecting according to a moderator-driven selection process in which a "moderator" in the environment is able to control the talking and listening status of the other users, and/or

selecting based on the geography or topology of the virtual environment, in which features of the environment are arranged realistically to affect the listening capability of users in the environment, so as to provide a coherent visual and sonic landscapes.

40. (Previously presented) A computer-readable medium according to claim 34, wherein the method further comprises:

generating for the particular room or area in the virtual user environment an in-room signal representative of the background due to audio streams emanating from users within the particular room or area and optionally including a weighted reverberant version of the adjoining room signal emanating from users in any room adjoining the particular room or area, wherein said generating includes combining the audio streams emanating from users within the particular room or area and optionally a weighted reverberant version of the combined signals emanating from users in the room or rooms or area or areas adjoining the particular room or area.

41. (Previously presented) A computer-readable medium according to claim 34, wherein the combining the audio streams in adjoining rooms or areas includes generating for the particular room or area a room signal representative of a reverberant version of combination of the signals emanating from users in the room adjoining the particular room or area.
42. (Previously presented) A method according to claim 36, wherein the adjoining room signal is formed by a process that includes combining the audio streams in adjoining rooms or areas including generating for the particular room or area a room signal representative of a reverberant version of combination of the signals emanating from users in the room adjoining the particular room or area.
43. (Previously presented) A computer readable medium according to claim 38, , wherein the virtual environment has one or more barriers defining two or more rooms or areas, wherein the particular user is located in the virtual environment in a particular room or area that includes an adjoining room or area that has an entrance or exit in the virtual user environment, wherein the rendering system is further configured to accept an adjoining room signal located at the entrance or exit, and wherein the adjoining room signal is representative of the combined noise emanating from the room or area adjoining the particular room or area at the entrance or exit.
44. (Currently amended) A computer readable medium according to claim 43, wherein the method further comprises:
- receiving an in-room signal representative of the background due to audio streams emanating from users within the particular room or area and optionally including a weighted reverberant versions of any adjoining room signal emanating from users in any room or area adjoining the particular room or area, wherein the in-room signal is formed by a process that includes combining the audio streams emanating from users within the particular room or area and optionally a weighted reverberant version of the combined signals emanating from users in the room or area or rooms or ~~rooms~~ areas adjoining the particular room or area.

45. (Previously presented) A computer readable medium according to claim 43, wherein the adjoining room signal is formed by a process that includes combining the audio streams in adjoining rooms or areas including generating for the particular room or area a room signal representative of a reverberant version of combination of the signals emanating from users in the room adjoining the particular room or area.